Referee response

Please see the detailed response to the referee reports below, followed by a redlined version of the manuscript with marked changes compared to the previous submission.

Response to report 1

We thank the referee for the time and effort in reviewing our manuscript. Below we address the conceptual and technical questions posed in their report.

> Recently, the physics of non-Hermitian systems has attracted considerable interest in both theory and experiments. The role of disorder and concomitant localization transitions in non-Hermitian systems has also been actively studied. In this context, I believe that this manuscript, which finds a new type of dynamical phase transitions in non-Hermitian disordered systems, should make a significant contribution in non-Hermitian physics.

We thank the referee for their positive assessment.

> (I) The authors claim that the obtained critical exponent for the current and drift velocity is 1/2. For example, the abstract reads, "The critical exponent of the transition equals 1/2 in propagating-propagating transitions". Additionally, in Sec. 5, the manuscript reads, "At the transition between distinct propagating phases, we found a $\nu = 1/2$ critical exponent. At the localization transition, the scaling also approaches $\nu = 1/2$". However, all the numerically obtained critical exponents summarized in Table I significantly deviate from $\nu = 1/2$. Similarly, in the last paragraph of Sec. 3, the manuscript reads, "Although we have no analytical argument for the scaling of v_{drift}, it also approaches to follow $\nu = 1/2$ scaling". However, I do not believe that the authors' numerical results approach $\nu = 1/2$ within the error bars, and do not find these explanations convincing. On the basis of the present numerical results, I believe that the statement that the "The critical exponent of the transition equals 1/2" is scientifically wrong. The authors should clarify this point.
>
> I suspect that the authors' expectation of $\nu = 1/2$ is due to the implicit assumption that this critical exponent should be universal (i.e., do not depend on specific details of systems but solely on fundamental properties such as symmetry and dimension) and be simple in one dimension. However, the universality of the authors' critical exponent ν is unclear only from the present results, in contrast to the conventional critical exponents that universally characterize the Anderson transitions (please see also below). Thus, I cannot rule out the possibility that the authors' critical exponent is different even in the same symmetry class and spatial dimension, and is not a simple rational number even in one dimension; I believe that it is significant to make a correct statement, rather than an intuitive but naive speculation.

We thank the referee for this feedback, and we agree that the discussion of our results was not sufficiently precise. In the updated manuscript we clearly state the observed difference between the numerical results and the scaling estimate.

While we cannot rule out that the critical exponent is non-universal, we would like to point out that the analytical argument for the metal-metal transition is only based on the presence of two maxima of $\mathrm{Im}(E)$ in the complex plane, rather than any specific details of the model. In our opinion this is a strong argument in favor of the universality of the critical exponent.

> (II) I fail to clearly understand how general and universal the authors' results are, in contrast to the conventional quantities whose critical exponents universally characterize the Anderson transitions (e.g., localization length, conductance). In particular, it is unclear whether we have a single critical exponent in the same symmetry class and spatial dimension. My concern here is partially because the authors' phase transition of the wave packet dynamics depends only on a few complex-valued eigenvalues that have the largest imaginary part, which contrasts with the conventional Anderson transitions that arise from the collective behavior of quasiparticles. For example, in the last paragraph of Sec. 4, the manuscript reads, "Here $\mathrm{Re}(J)$ does not exhibit finite-size scaling and therefore does not show a phase transition". This implies that the nature of phase transitions depends on specific details of models even in the same symmetry class and spatial dimension. Furthermore, in the first paragraph of Sec. 5, the manuscript reads, "Focusing on the metal-metal transition, we presented an analytical argument that proves that the value of 1/2 is universal". Although I agree that the authors provide an analytical argument, it still assumes some details of models that are not specified solely by symmetry and dimension, and I do not find the argument fully convincing. While the comprehensive discussions may go beyond the aim and scope of the present manuscript, I would like to request that the authors elaborate more on the universality of their results.

We agree with the referee that the exact degree of the universality of the wave packet localization transition remains an open question. However, as stated above, the analytical argument for the metal-metal transition critical exponent is a strong argument in favor of at least partial universality. The lack of scaling of $\mathrm{Re}(J)$ at the metal-insulator transition, however, is evidence that the metal-insulator wave packet transition differs from the metal-metal one.

> (III) While the authors' results [e.g., Fig. 1(b)] clearly show a phase transition as a consequence of the competition between non-Hermiticity and disorder, I fail to clearly understand whether this phase transition is continuous or discontinuous. Correspondingly, while the authors numerically obtain the critical exponents, I cannot exclude the possibility that this phase transition is actually discontinuous and that the obtained critical exponents are due to the finite-size effect and not well defined in the infinite-size limit. While some of the authors' results may already support the continuous phase transitions, I cannot clearly understand it in the present manuscript. Thus, I would like to request that the authors clearly demonstrate that the phase transitions observed in this manuscript are indeed continuous and are characterized by well-defined critical exponents.

Our manuscript demonstrates that the wave packet transitions are discontinuous in the infinite size limit and the critical exponents describe the finite-size scaling of the transition. This is demonstrated by the width of the transition vanishing in the infinite size limit (see Fig. 2), and is similar to the Anderson transitions in various symmetry classes, where the conductivity or Hall conductance has a discontinuity in the infinite size limit.

> (i) In the third paragraph of Sec. 2, the initial velocity is specified by "$k_x = \pi/2, k_y = 0$". I find this condition unclear simply because the two momenta k_x and k_y are given even in one dimension. The authors should clarify this point. While k_y may be relevant to the calculations of the two-dimensional model in Appendix E, it should not be written here since it is confusing.

We thank the referee for finding this oversight, and we have removed the mention of k_y from the 1D model.

> (ii) In Eq. (2), the authors provide the equation of motion for the wave packet dynamics in non-Hermitian Hamiltonians. I find some of the descriptions ambiguous and unclear. First, while the drift velocity is defined as "$\partial_t \langle\psi|x|\psi\rangle/\langle\psi|/\psi\rangle$", it is unclear whether this means "$\partial_t[\langle\psi|x|\psi\rangle]/\langle\psi|/\psi\rangle$" or $\partial_t[\langle\psi|x|\psi\rangle/\langle\psi|/\psi\rangle]$" (I suspect that the latter seems reasonable, though). Additionally, while the manuscript reads "... where we normalize the wave function such that $\langle\psi|\psi\rangle = 1$" just after Eq. (2), it is unclear when this normalization is introduced since it does not seem to be introduced before Eq. (2). Since the norm of wave functions is time dependent for non-Hermitian Hamiltonians, the precise way of normalization affects the way of the time derivative ∂_t and should be important. Correspondingly, I think that it would be better to provide a more detailed explanation on the derivation of Eq. (2).

We agree that this expression requires brackets, and further explanation of how this expression is obtained makes clear the assumptions that have been made. We added brackets over the entire expression $\langle\psi|x|\psi\rangle/\langle\psi|\psi\rangle$. We also provided a more detailed explanation of the derivation of Eq. (2) in a new Appendix.

> (iii) In the fifth paragraph of Sec. 2, the current for non-Hermitian Hamiltonians is introduced. However, the validity of this definition is unclear. More specifically, the authors define the current operator J by $J = -\partial_k H$. While this definition is arguably reasonable for Hermitian Hamiltonians H, its validity for non-Hermitian Hamiltonians is unclear. Correspondingly, for non-Hermitian Hamiltonians H, this current operator can be non-Hermitian, as is indeed the case for Eq. (3). While the authors only focus on the real part of the non-Hermitian current operator, they clarify the meaning of the imaginary part.

We have defined the current operator using the time derivative of the center of mass position of the wave packet, see App. C. We believe this fully addresses possible concerns about the validity of its definition.

> (iv) In Sec. 3, I fail to clearly understand the precise meaning of "direct" in the terminology "direct transition", which the authors should clarify.

By direct transition we mean a transition between two metallic phases that are not separated by an extended region of an insulating phase. We, however, agree that this is an uncommon term and we have replaced it with metal-metal transition in the revised manuscript.

Response to report 2

We thank the referee for the time and effort in reviewing our manuscript. Below we address the technical and conceptual questions posed in their report.

> The authors studied the wave packet dynamics in random non-Hermitian systems, where they found localization-unidirectional amplification transition and transition between propagating phases. I found the theme of the paper interesting, but the readers of this paper will not be able to reproduce the results, since some of the information for the numerical calculations are not provided in the paper.

We thank the referee for considering the reproducibility of our results. In order to ensure reproducibility of the data and figures shown in the paper, we provide a code repository as cited in the data availability section, located after the conclusion section. Further, all of the parameter values required to reproduce our results are listed in Appendix A.

> I also have several comments on the results.
>
> 1. One of the important finding of this paper is the scaling behaviors of v_{drift} and $\Re(J)$ (Figs. 2 (b)(c) and 3(c)). But I don't see the scaling equation in the text. The physical quantity such as v_{drift} is a function of several system parameters like L, δ_0, δ_1,...,h, phi, and the scaling form should be v_{drift}/(a W)=f(a, L, δ_0, δ_1,...,h, phi,..)= F(x(a, δ_0,...)$L^{1/\nu}$) where x is the relevant scaling variable. Is this what the authors mean? What is the relation between this x and b in the paper?

Our scaling equation listed in the text for the H_8 model, have $a(x)\tanh(b(x)\phi)$, with $x = L$ for $\Re(J)$ fits and t_{max} for v_{drift}. This is stated in the last paragraph of Section III, and we have now made it a numbered equation so that it is easier to locate. We outline the fit method of the transition width of the localization transition of the Hatano-Nelson model in the second paragraph of Section IV.

> 2. The scaling exponent nu=1/2 is outside the confidence intervals in Table 1. Do the authors think the exponent is 1/2 as they write in the abstract and Conclusion? If it is the case, the authors should explain why 1/2 is outside the confidence intervals of Table 1.

We agree with the referee that the numerical results for the exponents deviate significantly from $\nu = 1/2$, as the errors are too small to compensate for the discrepancy, and this warrants modifying our claims. We have now revised our statements to the following. Rather than claiming that we have found $\nu = 1/2$ numerically, we now state that finite-size effects may play an important role in the deviation from $\nu = 1/2$, that make the deviation larger than statistical errors can compensate for. Even after modifying these statements, what we still show with our work is that $\nu \neq 1$, which deviates from the single-energy expectation.

> 3. Below Eq. (1), the authors write that they set the initial velocity as $k_x = \pi/2$ and $k_y = 0$. What is k_y? I think the authors are considering one dimensional system.

We thank the referee for identifying this oversight. We have relegated the mention of k_y to Appendix E, which contains the discussion of two dimensional systems.

> 4. In the caption of Fig. 1, the value of delta is written as 0.01 and 0.03. What is this delta? $delta_i$ (i=0,1,2,3 and 4) are defined in Eqs. (1) and (4), but not delta.

We thank the referee for bringing this mistake to our attention. We have modified the Figure captions to include definitions of δ in terms of the δ_k as defined in the text, which are the standard deviation of the normal distribution from which disorder is sampled. Since the δ_k depend on the model, each Figure specifies which δ_k were used and their values, as well as the value of δ in terms of these δ_k.

> 5. I also think the values of h characterizing the asymmetry of hopping should be given.

We thank the referee for pointing out the missing value of h in the main text and Figure captions. The values of h are listed in Appendix A, along with other model and plotting parameters.

> 6. The authors write that h fixes the degree of non-Hermiticity. But I think the difference between $U_{1,j}$ and $U_{2,j}$ is also the origin of non-Hermiticity.

We agree with the referee that the non-Hermitian disorder terms $U_{i,j}$ contribute to the non-Hermiticity of the system. We are now more specific and say that h dictates the hopping asymmetry.

> 7. The authors used Taylor expansion to follow the dynamics. Isn't Chebyshev expansion better? Or is there a difficulty in applying Chebyshev expansion for non-Hermitian systems? If it is the latter case, the authors should comment on this.

We use Taylor expansions instead of Chebyshev expansions because they are the simplest to use. We have checked the validity of its application to our systems by studying the errors and verifying that they remain low.

> 8. In the definition of the drift velocity, $\partial_t \langle\psi|x|\psi\rangle / \langle\psi|/psi\rangle$, the authors should clarify if the time derivative operates on the normalization factor $\langle\psi|\psi\rangle$ or not.

We thank the referee for pointing out an ambiguous notation, which we now fixed. We also provide an extended derivation of the expression for the current in the new App. C.

> 9. I have difficulty understanding Fig. 2(a), which the authors call "real-space spectra". Are the axes real and imaginary parts of eigen energy? What is the definition of epsilon and E in this figure?

We thank the referee for pointing out that the definitions of E and ϵ are missing from the Figure 2 caption. We define E and ϵ in the first paragraph of Section II.

> 10. Why do we call the Hamiltonian in Eq. (4) as H_8?

H_8 was chosen because it reflects the shape of the PBC spectrum of the Hamiltonian, a figure-eight shape. We have now changed the notation to H_∞ in order to match the shape of the spectrum shown in Fig. 2 (a), and stated the reason for the notation in the main text.

Response to report 3

We thank the referee for the time and effort in reviewing our manuscript. Below we address the conceptual and technical questions posed in their report.

This paper describes a theoretical analysis of the localization transition in non-Hermitian lattices. Using a couple of specific models, it uncovers the processes by which the transition from a localizing to a non-localizing phase sets in for these lattices, which appears to be qualitatively different from the standard metal-to-insulator transition of Hermitian models. Essentially, any arbitrary initial wavepacket will evolve into the lattice's maximally amplified mode---the one with the largest value of Im(E)---so whether or not localization sets in depends on the dynamical characteristics of this maximally amplified mode. As disorder increases, it is possible for a mode with different characteristics to take over as the maximally amplified mode, giving rise to a localization transition.

One concern I have is that the introduction and conclusion refer to non-Hermitian systems as a general class, with the implication that these findings apply to other non-Hermitian systems, at least in 1D. Yet the study is based on the 1D Hatano-Nelson (HN) model, and other models with "point gaps" (spectra that form one or more loops in the complex energy plane), with periodic boundary conditions. As the authors are aware, the HN model and its variants have rather special properties not present in other non-Hermitian models. For one thing, the HN model's non-Hermitian hoppings bias propagation along one direction; for another, lattices with point gaps exhibit a "non-Hermitian skin effect" causing the periodic and finite lattice have extremely different eigenstates.

Thus, even if the conclusions of the study itself are sound, it doesn't seem warranted to confidently draw lessons for other kinds of non-Hermitian systems. Note also that the class of non-Hermitian lattices most easily realized in photonics or electrical circuits (as mentioned in the conclusion) is those with on-site loss and no point gaps.

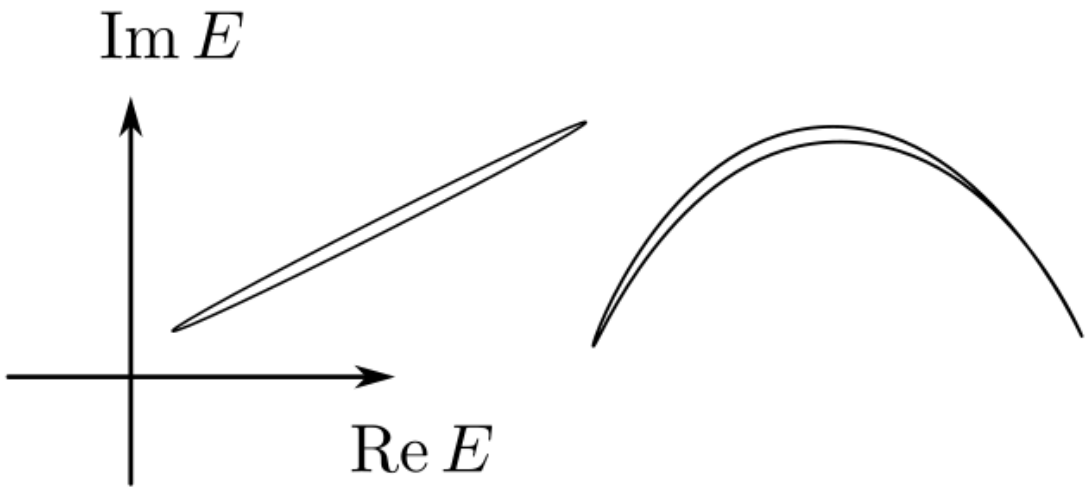
Within the context of the HN model and its relatives, I find it strange that the authors provide scarcely any physically-motivated discussion of the model's characteristics (e.g., the disorder-free HN model's propagation direction bias), and how they relate to the findings. For instance, in Fig. 1(c) it seems evident that the loop in the complex energy plane is continuable to the spectrum of the disorder-free HN model (roughly, waves moving along/opposite the bias direction get amplified/damped). With more disorder, the most affected eigenstates, whose energies migrate furthest away from the loop, are naturally those at the band extrema.

There are a couple of more minor issues with the figures, which will be detailed in the next section. If these requests can be satisfactorily accommodated, the paper can be accepted.

1. Either the introduction and conclusion should be toned down, or the authors should explain why their results ought to generalize.
2. Given the recent interest in point gaps, the non-Hermitian skin effect, etc., it would also be good if the authors could explicitly comment about whether their findings --- particularly the route to localization transition via the "maximally amplified wave packet" mechanism --- should hold for non-Hermitian systems that don't have point gaps, and/or those with open boundary conditions.
3. Consider providing more discussion of how the disorder-free HN model properties lead to the localization transition, as mentioned above.

We thank the referee for their remarks and start by addressing the generality of our results. We expect our results to hold generally because the scaling argument for the metal-metal transition is only based on the qualitative features of the spectrum, rather than any specific details of the model. We have added a corresponding statement to the argument itself and the conclusion.

The presence of a maximally amplified wave packet is generic to any disordered non-Hermitian system, since any non-Hermitian model will have a complex spectrum, and eigenstates at the top of the band dominate the dynamics, independent of the nature of the gap, the number of orbitals, or dimension. However, this maximally amplified wave packet will only have a well-defined unidirectional motion in systems with point gaps. To illustrate this, consider two possible dispersions without point gap on the complex plane sketched below.



The dispersion relation on the left has its maximally amplified eigenstate at a point of $v_{\mathrm{drift}} = 0$, while the dispersion relation on the right has a doubly degenerate maximally amplified eigenstate with opposite velocities. Therefore we expect that neither case exhibits unidirectional motion of the maximally amplified wave packet and that point gaps are necessary to observe maximally amplified wave packets that have a well-defined, finite drift velocity. We have added a corresponding statement to the discussion of the dispersion of Fig. 1c.

Because the phase transition is defined in the infinite size limit, in finite system with open boundary conditions it only describes transient dynamics. Our findings are not sensitive to the boundary conditions of the system the wave packets evolve in, as long as the system converges to the maximally amplified eigenstate before encountering the edges of the system.

> 4. Fix the following issues with the figures:
>
> - Below Eq. (1), the authors define multiple disorder parameters, denoted by δ_k. However, Fig. 1 refers to a single disorder parameter, delta, with no accompanying explanation. The other figures have a similar issue.

We thank the referee for bringing this mistake to our attention. We have modified the Figure captions to include definitions of δ in terms of the δ_k as defined in the text, which are the standard deviation of the normal distribution from which disorder is sampled. Since the δ_k depend on the model, each Figure specifies which δ_k were used and their values, as well as the value of δ in terms of these δ_k.

> - In Fig. 1(c)–(d), the axes should include ticks for at least the zero point.

We thank the referee for his remark and have modified the Figure as suggested. We have also added the zero points to Fig. 3 (a)-(b).

> - In Fig. 1, it would be nice if the authors could "close the loop" of the argument by showing that the largest-ϵ state indeed produces the time-domain results shown in (a).

In a finite system, a general starting vector in the long time limit converging to the maximally amplified eigenstate follows from the Schrödinger equation written in its eigenbasis. We therefore do not believe that it is necessary to show this explicitly in the text.

Phase transitions of wave packet dynamics in disordered non-Hermitian systems

Hélène Spring[1*], Viktor Könye[2], Fabian A. Gerritsma[1], Ion Cosma Fulga[2], and Anton R. Akhmerov[1]

1 Kavli Institute of Nanoscience, Delft University of Technology, P.O. Box 4056, 2600 GA Delft, The Netherlands
2 Institute for Theoretical Solid State Physics, IFW Dresden and Würzburg-Dresden Cluster of Excellence ct.qmat, Helmholtzstr. 20, 01069 Dresden, Germany
*helene.spring@outlook.com

Abstract

Disorder can localize the eigenstates of one-dimensional non-Hermitian systems, leading to an Anderson transition with a critical exponent of 1. We show that, due to the lack of energy conservation, the dynamics of individual, real-space wave packets follows a different behavior. Both transitions between localization and unidirectional amplification, as well as transitions between distinct propagating phases become possible. The critical exponent of the transition ~~equals~~ is close to $1/2$ in propagating-propagating and (de)localization transitions.

1 Introduction

Wave propagation in a strongly disordered medium stops due to Anderson localization [1]. The latter depends only on macroscopic properties of the medium, such as its dimensionality, symmetries, and topological invariants. In one space dimension (1D), for instance, generic disorder will localize all eigenstates, even if the disorder strength is infinitesimally weak. On the other hand, weak anti-localization becomes possible in two- and higher-dimensional systems, depending on their symmetries [2]. In such cases, the full spectrum of a disordered energy-conserving medium contains regions of localized and extended states, which are separated by mobility edges.

Unlike energy-conserving media, non-Hermitian systems with point gaps can exhibit fundamentally different behaviors in the presence of disorder. For instance, in the absence of energy conservation, it was found that weak disorder does not localize all states, even in 1D systems ~~[3–5]~~ [3,5]. Instead, similar to their higher-dimensional Hermitian counterparts, in 1D non-Hermitian systems localized and delocalized eigenstates are separated by mobility edges across which the localization length diverges. A recent work has shown that this divergence is governed by a universal critical exponent taking the value $\nu = 1$ [6].

One of the practical uses of the theory of eigenstate localization is to predict the dynamics of individual wave packets. In Hermitian systems, this is straightforward: the initial wave packet is decomposed into a superposition of states with different energies. The wave packet components above the mobility edge diffuse through the medium, while those below the mobility edge stay localized. By contrast, non-Hermitian systems break energy conservation, such that it is no longer possible to directly describe the wave packet dynamics by separating it into components with different energies.

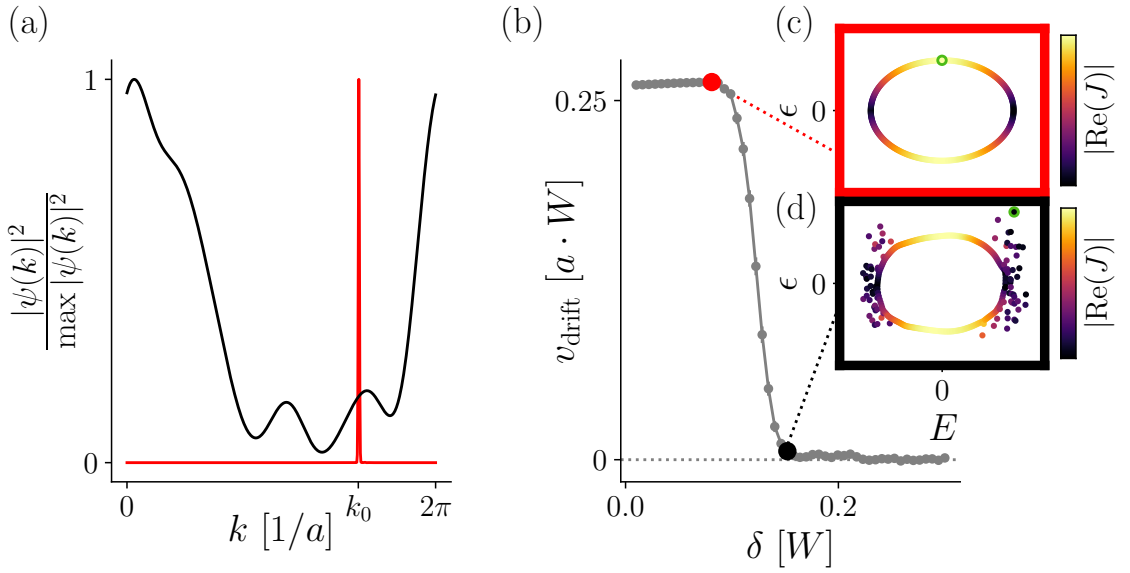


Figure 1: Maximally amplified waveforms of disordered Hatano-Nelson systems Eq. (1), with disorder ~~δ~~ in units of the model bandwidth W. Random disorder terms $U_{i,j}$ with $i \in \{0,1,2\}$ and j the site number (1) are sampled from distributions with standard deviation δ_i. We set $\delta_0 = \delta_1 = \delta_2 \equiv \delta$. (a) Magnitude of the Fourier components $|\psi(k)|^2$ of a wave packet evolved under H_{HN} for $\delta = 0.01$ (red) and $\delta = 0.3$ (black). The maximally amplified Fourier component of the system with low disorder is marked by k_0. (b) The average drift velocity v_{drift} as a function of disorder strength δ, and a the lattice constant. (c) Eigenvalues of H_{HN} for a single disorder realization with disorder strength $\delta = 0.08$. The point of maximal amplification ϵ_{max} is highlighted with green. (d) Eigenvalues of a disordered system with disorder $\delta = 0.15$. ϵ_{max} is highlighted in green. Plot details in App. A.

Here we demonstrate that the difference between single energies and wave packets is profound. Because in the long-time limit any wave packet in a system with a point gap converges to a maximally amplified waveform, the asymptotic shape of the wave packet may change discontinuously when the system parameters are varied. This enables a ~~direct~~ metal-metal transition between different unidirectionally amplified phases in addition to the previously known localization transition. Furthermore, in finite-size systems the fluctuations of the maximally amplified energy are self-averaging, which results in a critical exponent ~~of $\nu = 1/2$~~$\nu \approx 1/2$ that is close to our analytical estimate of 1/2.

The structure of the manuscript is as follows. In Sec. 2 we demonstrate the universal convergence of wave packets in weakly disordered systems to the maximally amplified waveform. In Sec. 3 we study the ~~direct~~ transition between distinct propagating phases. In Sec. 4 we show that the wave packet single-frequency transition differs from the static non-Hermitian single-frequency transition. We conclude in Sec. 5.

2 Maximally amplified wave packet

Unlike their Hermitian counterparts, one-dimensional (1D) non-Hermitian systems with no symmetries do not localize in the presence of weak disorder [3,7,8]. The different Fourier components of the wave packet, which are coupled by scattering events, are amplified at

different rates, depending on the value of ϵ, the imaginary part of their energy $E + i\epsilon$ [5]. The eigenstate whose eigenvalue has the largest positive imaginary component, $\epsilon_{\max}$, is amplified the fastest. This means that any waveform in a weakly disordered medium converges to the maximally amplified waveform, forming an envelope in Fourier space around the point of maximal amplification k_0.

To demonstrate this we consider a Hatano-Nelson Hamiltonian [3]:

$$H_{\text{HN}} = \sum_j U_{0,j}|j\rangle\langle j| + \left(-\frac{W}{2}e^{-h} + U_{1,j}\right)|j\rangle\langle j+1| + \left(-\frac{W}{2}e^{h} + U_{2,j}\right)|j+1\rangle\langle j|, \quad (1)$$

where the sum runs over sites j of the system, W is a hopping parameter that sets the bandwidth of the system, h fixes the ~~degree of non-Hermiticity~~magnitude of the hopping asymmetry, and $U_{k,j}$ are the complex disorder coefficients whose real and imaginary parts are independently sampled from a normal distribution with zero mean and standard deviation δ_k. Thus, δ_k models the strength of each type of disorder (onsite or hopping).

We time-evolve wave packets numerically by Taylor expanding the time-dependent Schrödinger equation to first order [See App. B for numerical methods]. For concreteness, throughout the following we consider an initial wave packet that has a Gaussian profile $u(x) = e^{-ikx}e^{-(x-x_0)/2\sigma^2}$. This wave packet is initialized at the center of the periodically wrapped lattice ($x_0 = 0$), with a width one tenth of the width of the lattice ($\sigma = L/10$) and with the same initial velocity (~~$k_x = \pi/2,\ k_y = 0$~~$k_x = \pi/2$) for all simulations. The wave packet evolving under the weakly disordered Hatano-Nelson model Eq. (1) converges to an envelope around the point of maximal amplification k_0 [Fig. 1 (a), red curve], where k_0 is the k-point corresponding to the eigenvalue with the largest positive imaginary part, calculated from the dispersion of the Hamiltonian without disorder. For large disorder, the waveform acquires a non-universal shape whose center of mass is not guaranteed to be located around k_0 [Fig. 1 (a), black curve] .

The motion of the center of mass of the waveform in real space defines the drift velocity of the wave packet, ~~$v_{\text{drift}} = \partial_t\langle\psi|\hat{x}|\psi\rangle/\langle\psi|\psi\rangle$~~$v_{\text{drift}} = \partial_t[\langle\psi|\hat{x}|\psi\rangle/\langle\psi|\psi\rangle]$, with $\hat{x}$ the position operator. We evaluate this expression and obtain:

$$\partial_t\langle\psi|\hat{x}|\psi\rangle/\langle\psi|\psi\rangle = \frac{1}{2}\langle\psi|\partial_k\left(H + H^\dagger\right)|\psi\rangle + \frac{i}{2}\langle\psi|\{H - H^\dagger, \hat{x} - \langle\psi|\hat{x}|\psi\rangle\}|\psi\rangle,$$

$$\partial_t[\langle\psi|\hat{x}|\psi\rangle/\langle\psi|\psi\rangle] = \frac{1}{2}\langle\psi|\partial_k\left(H + H^\dagger\right)|\psi\rangle + \frac{i}{2}\langle\psi|\{H - H^\dagger, \langle\psi|\hat{x}|\psi\rangle - \hat{x}\}|\psi\rangle, \quad (2)$$

where $\{\cdot,\cdot\}$ is the ~~anti-commutator~~ anticommutator and where we normalize the wave function such that $\langle\psi|\psi\rangle = 1$. Details of the derivation of Eq. (2) are in Appendix C.

The momentum-space non-Hermitian generalization of the current associated with a Hamiltonian H is defined as $J(H) = -\partial_k H$. The first term of (2) is $\text{Re}(\langle\psi|J|\psi\rangle)$ and for a single Bloch state k_0, $\partial_t\langle\psi|\hat{x}|\psi\rangle_{k_0} = \text{Re}(J)|_{k_0}$. For the Hatano-Nelson Hamiltonian (1),

$$J(H_{\text{HN}}) = \sum_j i\left(-\frac{W}{2}e^{-h} + U_{1,j}\right)|j\rangle\langle j+1| + i\left(\frac{W}{2}e^{h} - U_{2,j}\right)|j+1\rangle\langle j|.$$

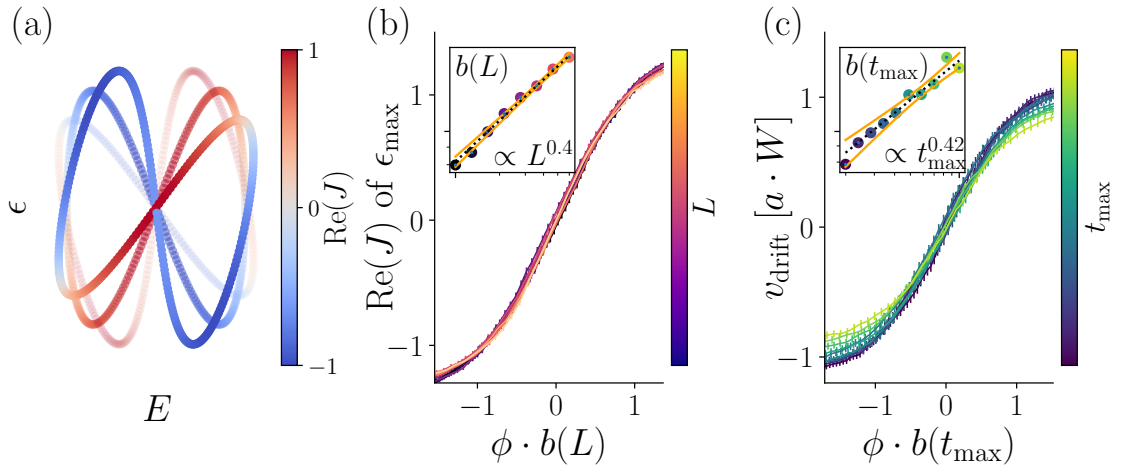


Figure 2: Phase transition between left and right moving wave packets of Hamiltonian ~~H_8 Eq.~~ H_∞ (4), with onsite and hopping disorder strength ~~$\delta = 0.1$~~ in units of the system bandwidth W. Random disorder terms $U_{i,j}$ with $i \in \{0, 1, 2, 3, 4\}$ and j the site number (4) are sampled from distributions with standard deviation δ_i. We set $\delta_0 = \delta_1 = \delta_2 = \delta_3 = \delta_4 \equiv \delta = 0.1$. (a) Real-space spectra for $\phi = 0.3$ (most transparent), $\phi = 0$ (intermediate transparency) and $\phi = -0.3$ (most opaque). (b) Rescaled Re(J) of the maximally amplified eigenstate and (c) v_{drift} around $\phi = 0$. Insets: scaling function of the slope at the transition and the 95% confidence interval. Plot details in App. A.

$$J(H_{\text{HN}}) = \sum_j i\left(-\frac{W}{2}e^{-h} + U_{1,j}\right)|j\rangle\langle j+1| + i\left(\frac{W}{2}e^{h} - U_{2,j}\right)|j+1\rangle\langle j|. \tag{3}$$

At the localization transition, the drift velocity of the wave packet v_{drift} falls to 0 [Fig. 1 (b)]. We observe that below the localization transition, v_{drift} is finite and Re(J) at $\epsilon_{\max}$ is also finite [Fig. 1 (c)], and likewise when the wave packet is localized the Re(J) at $\epsilon_{\max}$ is 0 [Fig. 1 (d)]. Finiteness of the drift velocity and Re(J) at $\epsilon_{\max}$ are a consequence of the existence of a point gap in the spectrum of the Hamiltonian: a finite area enclosed by the dispersion relation of the Hamiltonian in the complex plane.

Disorder shifts eigenvalues around in the complex plane, resulting in a different eigenstate becoming maximally amplified. Disorder also nontrivially changes the Re(J) of these eigenvalues. For strong disorder, the maximally amplified eigenstate is generically localized and Re(J) = 0. The maximally amplified state may have nonzero Re(J) [Fig. 1 (c)], and therefore be delocalized [Fig. 1 (b)] or have zero Re(J) [Fig. 1 (d)], and therefore be localized [Fig. 1 (b)]. If that state is delocalized, then the system delocalizes. Likewise if it is localized the system is localized even if other states in the systems are delocalized, since these states are always less amplified than the state at $\epsilon_{\max}$. Fig. 1 (d) shows that although delocalized states exist for $\epsilon < \epsilon_{\max}$, the system is localized because the maximally amplified state at $\epsilon_{\max}$ has Re(J) = 0.

3 ~~Direct transition~~ Transition between propagating phases

The expectation that propagating waveforms in non-Hermitian systems always evolve to the maximally amplified waveform suggests that a ~~direct~~ transition between competing propagating phases whose ϵ are close to $\epsilon_{\max}$ should be possible. Here we construct a

Model	Quantity	Scaling exponent
~~H_8~~ H_{∞} (4)	$\mathrm{Re}(J)$	0.41 ± 0.01
	v_{drift}	0.45 ± 0.03
H_{HN} (1)	v_{drift}	0.38 ± 0.04

Table 1: Scaling parameters of the phase transitions shown in Fig. 2 and 3.

Hamiltonian ~~that hosts states propagating with opposite velocities at an ϵ close to $\epsilon_{\max}$:~~ with an ∞-shaped dispersion relation that allows to discontinuously switch the velocity of a maximally localized wave vector as a function of its parameters [see Fig. 2 (a)]:

$$H_{\underline{8}\infty} = \sum_j U_{0,j}|j\rangle\langle j| + \left(\frac{We^{i\phi}}{2}e^{-h} + U_{1,j}\right)|j\rangle\langle j+1| + \left(\frac{We^{i\phi}}{2}e^{h} + U_{2,j}\right)|j+1\rangle\langle j| + \left(\frac{We^{i\phi}}{2}e^{-h} + U_{3,j}\right)|j\rangle\langle j+2| + \left(-\frac{We^{i\phi}}{2}e^{h} + U_{4,j}\right)|j+2\rangle\langle j|, \tag{4}$$

where the sum runs over sites j of the system, W is a hopping parameter that sets the bandwidth of the system, ϕ rotates the spectrum in the complex plane, and where $\mathrm{std}(U_{k,j}) = \delta_k$ as in (1). The non-Hermitian generalization of the current J is given by

$$J(H_{\underline{8}\infty}) = \sum_j i\left(\frac{We^{i\phi}}{2}e^{-h} + U_{1,j}\right)|j\rangle\langle j+1| - i\left(\frac{We^{i\phi}}{2}e^{h} + U_{2,j}\right)|j+1\rangle\langle j| + 2i\left(\frac{We^{i\phi}}{2}e^{-h} + U_{3,j}\right)|j\rangle\langle j+2| - 2i\left(\frac{We^{-i\phi}}{2}e^{h} + U_{4,j}\right)|j+2\rangle\langle j|. \tag{5}$$

The ~~spectrum of H_8 is composed of two lobes Fig. 2 (a). The~~ eigenstates associated to the eigenvalues at the top of the left lobe of the complex dispersion relation propagate to the left, ~~and likewise~~ while those at the top of the right lobe propagate right, as shown by the sign of $\mathrm{Re}(J)$. By continuously tuning ϕ through 0, there is a discontinuous change in the eigenvalue with the largest positive imaginary component [Fig. 2 (a)] which leads to an abrupt transition between two different maximally amplified eigenstates. When $\phi \neq 0$, wave packets are amplified either predominantly to the left or to the right. The maximally amplified eigenstate of ~~H_8~~ H_{∞} at $\phi = 0^-$ propagates to the left, and the one at $\phi = 0^+$ propagates to the right, meaning there is a metal-metal transition at $\phi = 0$. This transition is marked by a switch in the signs of both $\mathrm{Re}(J)$ and v_{drift} [Fig. 2 (b)-(c)].

In the presence of disorder and for finite system size, the average of $\mathrm{Re}(J)$ at $\epsilon_{\max}$ and v_{drift} changes linearly in the vicinity of $\phi = 0$, with an intermediate localized point at the middle of the transition. The slope of this transition increases with system size L (for $\mathrm{Re}(J)$), and the total number of simulated time steps $t_{\max}$ (for v_{drift}). We therefore confirm that the transition between the two propagating phases on either side of $\phi = 0$ does not go through a localized phase.

We examine finite-size scaling of the system at the transition. Due to the shape of the spectrum of ~~H_8~~ H_{∞} [Fig. 2 (a)] on either side of the transition, the distribution of E is bimodal, grouped around two values where ϵ is the largest. The variance of the individual peaks is the same at the transition point $\phi = 0$. Their standard deviations dictate the width of the transition, as $\phi \cdot t$ is required to be larger than these standard deviations

in order for one part of the spectrum, and therefore one value of Re(J) to 'win' over the other.

There are several considerations we can make in order to estimate the scaling of these standard deviations as a function of system size. The variance of the ~~peaks is equivalent~~ peak positions equals to the variance of the expectation value of the disorder $U(x)$ in the system, $\mathrm{var}(\langle\psi|\, U(x)\, |\psi\rangle) = \mathrm{var}(\int_0^L \psi^*(x)U(x)\psi(x)dx)$. We reach an analytical expression for the scaling of the variance of the peak by ~~considering that on either side of the transition, the system contains delocalized phases that behave like plane waves and propagate~~ assuming that the maximally localized wave packets spread throughout the system ~~. The modulus of these propagating waves is approximately constant,~~ on both sides of the transition so that $|\psi| \sim$ const. Therefore the dependence of the variance of the expectation value on system size L is given by $\mathrm{var}(L^{-1}\int_0^L U(x)dx) = L^{-2}\cdot \mathrm{var}\left(\int_0^L U(x)dx\right) \propto L^{-2}L = L^{-1}$. ~~The~~ From this we conclude that the standard deviation of each peak of the distribution of ϵ, and therefore the width of the transition, scales with $1/\sqrt{L}$. This leads to the expectation for the finite-size scaling of $b(L)$ to follow $\sqrt{L}$. ~~This is in direct contrast to the expectation from single-energy studies where~~ This estimate differs from the critical exponent ~~is~~ $\nu = 1$ [6] ~~. However, by construction the H_8 model transition is not a single-energy transition~~of the single energy delocalization transition and suggests that the two transitions are of different nature. This analytical argument relies solely the existence of two non-adjacent maxima of Im(E), and therefore applies to a broad class of models, suggesting that the metal-metal transition is universal.

~~We~~ To compute the critical exponent from the numerical simulations we fit v_{drift} and Re(J) of Fig. 2 with the function ~~$a\tanh(b\phi)$, where a, b~~

$$a\tanh(b\phi), \tag{6}$$

where a and b are functions of system size L for Re(J) fits, and functions of simulation time t_{max} for v_{drift}. We choose b as our relevant scaling parameter, since it measures the width of the transition. The fitting function $\tanh(b\phi)$ has a discontinuous jump at $L\to\infty$ similar to scaling variables in other systems, such as Hall conductivity σ_{xy} at quantum Hall plateau transitions. The numerical results for Re(J) at ϵ_{max} show that the ~~scaling is closer to $\nu = 1/2$ scaling than $\nu = 1$ scaling~~ critical exponent extracted from the scaling is $\nu \approx 1/2$, as predicted by our analytical estimate [~~see~~ inset of Fig. 2 (c)~~and~~]. At this moment we cannot deterimne whether the difference of the numerical results and the analytical estimate $\nu = 1/2$ as reported in Table 1 is potentially due to finite-size effects of the simulation or a missing aspect of the analytical estimate. Although we have no analytical argument for the scaling of v_{drift}, it also appears to ~~follow~~ be closer to $\nu = 1/2$ scaling than $\nu = 1$ scaling [see inset of Fig. 2 (d) and Table 1]. App. D contains further discussion of the bimodal behavior.

4 Metal-insulator transition

The metal-metal transition behaves differently from the single-frequency response, which raises the question whether the metal-insulator transition is also different. In the presence of non-Hermitian disorder in both the onsite and hopping terms, the metal-insulator transition of the Hatano-Nelson Hamiltonian is the result of a discontinuous change in ϵ_{max} [Fig. 1 (b)-(d)], and the same arguments as the metal-metal transition apply there. We therefore test whether a transition that does not involve a discontinuous switch of ϵ_{max} and E matches the single-frequency response. The original Hatano-Nelson Hamilto-

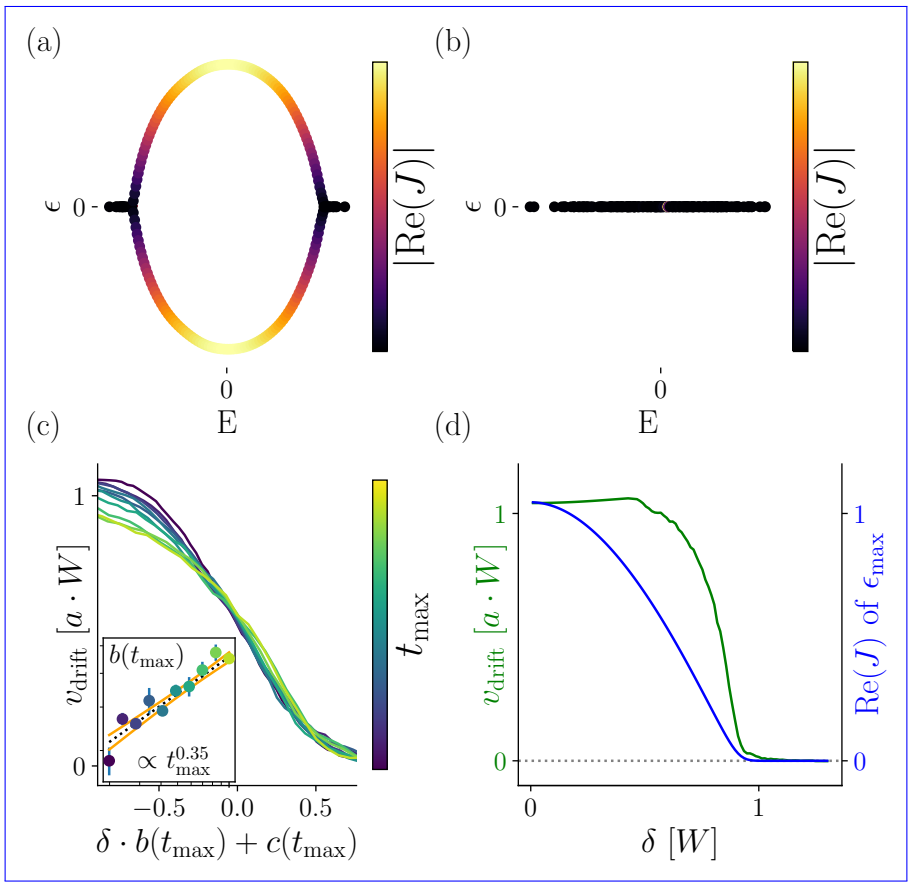


Figure 3: Finite-size scaling of the Hatano-Nelson Hamiltonian Eq. (1), with onsite disorder~~$\delta = \delta_0$~~. Random disorder terms $U_{i,j}$ with $i \in \{0, 1, 2\}$ and j the site number (1) are sampled from distributions with standard deviation δ_i. We set $\delta_1 = \delta_2 = 0$ and label δ_0 as δ. (a)-(b) The Hatano-Nelson spectrum for disorder strength (a) $\delta = 0.3$ and (b) $\delta = 1.2$. (c) Rescaled wave packet drift v_{drift} at the transition point, collapsed using the relevant scaling parameter $b(t_{\mathrm{max}})$ and the irrelevant scaling parameter $c(t_{\mathrm{max}})$. Inset: fit of the scaling parameter $b(t_{\mathrm{max}})$ and the 95% confidence interval. (d) Comparison of v_{drift} (green) to Re(J) of ϵ_{max} (blue) for system sizes $L = 10^3$. Plot details in App. A.

nian [3] fulfills this condition. We obtain this Hamiltonian by setting the disorder terms δ_i of Eq. (1) to be 0 except for δ_0. Here the maximally amplified state is the last state to localize, as the mobility edge moves from the largest absolute values of E to the smallest [Fig. 3 (a)-(b)].

The shapes of the $v_{\mathrm{drift}}(\delta)$ curves of Fig. 3 do not lend themselves to a tanh fit. The scaling variable b we choose in this case is the maximum slope during the transition. We also track an irrelevant scaling variable c to ensure the superposition of the rescaled curves. The v_{drift} curves do not fully collapse at the transition [Fig. 3 (c)]. The scaling of v_{drift} is $b(t_{\mathrm{max}}) \propto t_{\mathrm{max}}^{0.38}$. We have no analytical expectation for the scaling of v_{drift}, however we cannot rule out that the different critical exponent is due to finite-size effects and a finite resolution of the simulation, as demonstrated by the quality of the fit of the scaling parameter [see inset of Fig. 3 (c)]. ~~Regardless of the actual value, the~~ The scaling behavior differs from $\nu = 1$ and is close to $\nu = 1/2$.

Here Re(J) does not exhibit finite-size scaling and therefore does not show a phase transition. The phase transition of v_{drift} is thus ascribable to the non-linear term of Eq. (2) since it is absent from the linear term. Re(J) and v_{drift} nevertheless both fall to 0 at the

same point [Fig. 3 (d)]. When taking the biorthogonal expectation value to calculate $\mathrm{Re}(J)$, finite-size scaling does occur [See App. E for discussion].

5 Conclusion

We showed that the dominant dynamics are attributable to a single point in the Fourier space of wave packets, which corresponds to the maximally amplified eigenstate. In the long time limit and in the presence of disorder, wave packets follow a behavior that is independent of initial conditions because they converge to the maximally amplified waveform. ~~At the transition between distinct propagating phases, we found a $\nu = 1/2$ critical exponent . At the localization transition, the scaling also approaches~~ Focusing on the metal-metal transition, we presented an analytical argument that yields an expected value of the critical exponent of $\nu = 1/2$ ~~. This clearly proves~~ in any model with two separate maxima of the dispersion relation in the complex plane. For transitions between propagating phases as well as for localization transitions, our numerical results yield $\nu \approx 1/2$, with deviations from our analytical prediction $\nu = 1/2$ likely due to finite-size effects. The reasonable agreement between our numerics and our analytical estimate is a strong indication that $\nu = 1/2$ is universal. Our results clearly prove that wave packet transitions in disordered non-Hermitian media differ from the single-frequency response~~. Focusing on the metal-metal transition, we presented an analytical argument that proves that the value of 1/2 is universal~~, $\nu = 1$.

In our simulations we have observed that drift velocity of a wave packet $\mathrm{v}_{\mathrm{drift}}(t_{\mathrm{max}})$ follows a scaling law similar to the scaling of an eigenstate in a finite system. It is not obvious that this equivalence is guaranteed, and further studies are required.

The nature of transitions in higher-dimensional non-Hermitian systems remains an open question. Preliminary results for two-dimensional systems show that the critical exponent differs from both $\nu = 1$ and $\nu = 1/2$ [App. F]. It is therefore ~~possible~~ likely that the critical exponents of wave packet dynamics in non-Hermitian systems are dimension-dependent.

Non-Hermitian systems are naturally realizable in experiment, and non-Hermitian wave packet dynamics are studied in photonic lattices and electrical circuits [9–11, 11–15]. The ~~direct~~ transition between propagating phases can be implemented as a switch tuned by a continuous parameter, with uses in control or sensor systems.

Data and code availability

The data shown in the figures, as well as the code generating all of the data is available at [16].

Author contributions

The initial project idea was formulated by I. C. F. and was discussed and later refined with contributions from H. S., V. K., and A. A. Results and code for non-Hermitian spectra were produced by H. S. The initial version of the wave packet propagation code and the error estimation were produced by V. K. Wave packet propagation with disorder was simulated by F. G. under the supervision of H. S. A. A related to relevant literature and formulated the question of comparing wave packet and single frequency phase transitions. The final

data was generated by H. S. with input from I. C. F., A. A., and V. K. The scaling analysis was performed by H. S. with guidance from I. C. F. The manuscript was written by H. S. with contributions from I. C. F., A. A., and V. K. The project was managed by H. S. and A. A .

Acknowledgments

The authors thank D. Varjas and M. Wimmer for their contributions to formulating the project idea and for helpful discussions. The authors thank Ulrike Nitzsche for technical assistance. H. S. thanks A. L. R. Manesco for helpful discussions and for reviewing the project code. A. A. and H. S. were supported by NWO VIDI grant 016.Vidi.189.180 and by the Netherlands Organization for Scientific Research (NWO/OCW) as part of the Frontiers of Nanoscience program. I. C. F. and V. K. acknowledge financial support from the DFG through the Würzburg-Dresden Cluster of Excellence on Complexity and Topology in Quantum Matter – ct.qmat (EXC 2147, project-id 390858490).

References

[1] P. W. Anderson, *Absence of diffusion in certain random lattices*, Phys. Rev. **109**, 1492 (1958), doi:10.1103/PhysRev.109.1492.

[2] F. Evers and A. D. Mirlin, *Anderson transitions*, Rev. Mod. Phys. **80**, 1355 (2008), doi:10.1103/RevModPhys.80.1355.

[3] N. Hatano and D. R. Nelson, *Localization transitions in non-hermitian quantum mechanics*, Phys. Rev. Lett. **77**, 570 (1996), doi:10.1103/PhysRevLett.77.570.

[4] ~~N. Hatano, T. Watanabe and J. Yamasaki, *Localization, resonance and non-hermitian quantum mechanics*, Phys. A , 170 (2002), , Horizons in Complex Systems.~~

[5] P. W. Brouwer, P. G. Silvestrov and C. W. J. Beenakker, *Theory of directed localization in one dimension*, Phys. Rev. B **56**, R4333 (1997), doi:10.1103/PhysRevB.56.R4333.

[6] K. Kawabata and S. Ryu, *Nonunitary scaling theory of non-hermitian localization*, Phys. Rev. Lett. **126**, 166801 (2021), doi:10.1103/PhysRevLett.126.166801.

[7] K. Kawabata, M. Sato and K. Shiozaki, *Higher-order non-hermitian skin effect*, Phys. Rev. B **102**, 205118 (2020), doi:10.1103/PhysRevB.102.205118.

[8] H. Sahoo, R. Vijay and S. Mujumdar, *Anomalous transport regime in non-hermitian, anderson-localizing hybrid systems*, doi:10.48550/ARXIV.2206.05280 (2022).

[9] P. Zhu, X.-Q. Sun, T. L. Hughes and G. Bahl, *Higher rank chirality and non-hermitian skin effect in a topolectrical circuit*, doi:10.48550/ARXIV.2207.02228 (2022).

[10] M. Wu, Q. Zhao, L. Kang, M. Weng, Z. Chi, R. Peng, J. Liu, D. H. Werner, Y. Meng and J. Zhou, *Evidencing non-bloch dynamics in temporal topolectrical circuits*, doi:10.48550/ARXIV.2206.11542 (2022).

[11] Q. Lin, T. Li, L. Xiao, K. Wang, W. Yi and P. Xue, *Observation of non-hermitian topological anderson insulator in quantum dynamics*, Nat. Commun. **13** (2022), doi:10.1038/s41467-022-30938-9.

[12] S. O. Peatáin, T. Dixon, P. J. Meeson, J. Williams, S. Kafanov and Y. A. Pashkin, *The effect of parameter variations on the performance of the josephson travelling wave parametric amplifiers*, doi:10.48550/ARXIV.2112.07766 (2021).

[13] S. Weidemann, M. Kremer, S. Longhi and A. Szameit, *Non-hermitian anderson transport*, doi:10.48550/ARXIV.2007.00294 (2020).

[14] T. Jiang, A. Fang, Z.-Q. Zhang and C. T. Chan, *Anomalous anderson localization behavior in gain-loss balanced non-hermitian systems*, Nanophotonics **10**(1), 443 (2021), doi:doi:10.1515/nanoph-2020-0306.

[15] F. Noronha, J. A. S. Lourenço and T. Macrì, *Robust quantum boomerang effect in non-hermitian systems*, Phys. Rev. B **106**, 104310 (2022), doi:10.1103/PhysRevB.106.104310.

[16] H. Spring, V. Konye, F. A. Gerritsma, I. C. Fulga and A. R. Akhmerov, *Phase transitions of wave packet dynamics in disordered non-Hermitian systems*, doi:10.5281/zenodo.7535012 (2023).

[17] N. J. Higham, *Functions of Matrices: Theory and Computation*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, ISBN 978-0-898716-46-7 (2008).

[18] M. L. Liou, *A Novel Method of Evaluating Transient Response*, Proc. IEEE **54**(1), 20 (1966), doi:10.1109/PROC.1966.4569.

[19] C. Moler and C. Van Loan, *Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five Years Later*, SIAM Rev. **45**(1), 3 (2003), doi:10.1137/S00361445024180.

A ~~Numerical methods~~

~~The time evolution of the wave packets was calculated using the scaled Taylor expansion method to first order [17–19], obtaining~~

$$|\psi(t+dt)\rangle = |\psi(t)\rangle - iH|\psi(t)\rangle dt,$$

~~where $|\psi(t)\rangle$ is the wave function at time t, dt is the time step, and H is the Hamiltonian dictating the time evolution. The simulation time t and timesteps dt are in units of the bandwidth W of the system. We choose $dt = 0.01$, but we have separately checked that our results hold also for smaller time steps. In our simulations we initialize the system from the same real-space Gaussian wave packet. In order to ensure that the wave packets do not reach the system boundary, we limit the total number of time steps used for a simulation to $t_{\max} = L/(a \cdot dt \cdot v_{\text{drift}})$, with L the system size, a the lattice constant and v_{drift} the drift velocity of the wave packet for low disorder. Above the localization transition, $t_{\max}$ is not shortened in order to record instances of 'teleportation' of the drift center of the maximally amplified wave packet, which contribute to the average velocity.~~

~~The method is based on the following expression for the matrix exponential~~

$$\mathrm{e}^{-itH} = \lim_{N\to\infty} \left(I - \frac{itH}{N}\right)^N,$$

where N is the number of time steps. Fixing the time step ($dt = t/N$) and the number of steps (N) we get an approximation for the time evolution operator as:

$$e^{-itH} \approx (I - idtH)^N .$$

The error introduced at each subsequent time step can be estimated using the errors calculated for Taylor polynomials of the first order as [18, 19]:

$$\delta = \left\| e^{-idtH} - I + idtH \right\| \leq \frac{dt^2 \|H\|^2}{2} \frac{1}{1 - \frac{dt\|H\|}{3}},$$

where $\|\cdot\|$ is any well defined matrix norm, for simplicity we use the spectral norm. For normalized Hamiltonians $\|H\| = 1$ and $dt \leq 1$, the error introduced at each time step is $\delta \leq 3dt^2/4$.

A Model and plotting parameters

For Fig. 1 (b), δ is varied between 0.01 and 0.3 in 50 steps, and the average drift velocity is averaged over 600 different disorder configurations. The spectra of panels (c) and (d) are calculated for systems composed of 300 lattice sites, and parameter h set to 0.3. For panels (a) and (b), the wave packet evolution was performed on system sizes of 600 sites, in steps of $dt = 0.01$ for 60000 steps. For panel (a) the results displayed in the figure are taken at the last step of the time evolution. The disorder strength δ is given in units of W the bandwidth of H_{HN}.

For Fig. 2, the spectra, $\mathrm{Re}(J)$ and the wave packet results are obtained for systems with sizes $L \in \{199, 238, 285, 341, 408, 488, 584, 698, 836, 1000\}$. Results for $\mathrm{Re}(J)$ and wave packets are averaged over 2000 and 500 different disorder configurations respectively. The wave packet evolution was performed in steps of $dt = 0.01$ for L/dt steps. The tilt angle ϕ was varied between -0.1 and 0.1 in the following way: 20 points between -0.1 and -0.03, 100 points between -0.03 and 0.03, and 20 between 0.03 and 0.1. The disorder strength is set to $\delta = 0.1$ in units of the bandwidth W of the Hamiltonian Eq. (4).

For Fig. 3, the parameter h is set to 0.3. Ten different system sizes L are simulated, $L \in \{199, 238, 285, 341, 408, 488, 584, 698, 836, 1000\}$. Results are averaged over 500 different disorder configurations for each value of disorder strength. The wave packet evolution was performed in steps of $dt = 0.01$ for L/dt steps, with the values of L as stated above.

For the insets of Fig. 2 (b)-(c) and Fig. 3 (c), the error of the scaling fit is shown using the 95% confidence interval.

For Fig. 4, panel (a) data is made up of 5000 disorder configurations, for systems 800 sites long. Panel (c) data is made up of 3000 disorder configurations, for systems 800 sites long. Panel (c) data is made up of 2000 disorder configurations, for systems 1000 sites long. Panel (d) data is made up of 3000 disorder configurations, for systems 800 sites long and $\delta = 0.8$. The wave packet results are obtained for time evolution step size $dt = 0.01$ and total time steps L/dt.

Fig. 7 is composed of unscaled data that was obtained and used in Fig. 2 and Fig. 3.

For Fig. 6, the wave packet evolution was performed in steps of $dt = 0.01$ for L/dt steps. Five different system sizes $L \times L$ were simulated, with $L \in \{64, 85, 113, 150, 199\}$. Disorder strengths were varied between 0.01 and 0.5 in 50 steps. For each disorder strength, the result is averaged over 400 different disorder configurations.

For Fig. 5, panel (a) data for $\delta = 0.01$ is composed of 100 different disorder configurations for systems of 800 sites. Panel (a) data for $\delta = 0.1$ is made up of 1000 disorder

configurations, for systems 1000 sites long. Panel (b) data is made up of 5000 disorder configurations, for systems 800 sites long and $\delta = 0.8$. For panel (c), results for $\mathrm{Re}(J)$ and wave packets are averaged over 500 different disorder configurations. The tilt angle ϕ was varied between -0.1 and 0.1 in the following way: 20 points between -0.1 and -0.03, 100 points between -0.03 and 0.03, and 20 between 0.03 and 0.1. Results are obtained for systems with sizes $L \in \{199, 238, 285, 341, 408, 488, 584, 698, 836, 1000\}$. The disorder strength is set to $\delta = 0.1$ in units of the bandwidth W of the Hamiltonian (4). For the insets of panels (c)-(d), the error of the scaling fit is shown using the 95% confidence interval.

B Numerical methods

The time evolution of the wave packets was calculated using the scaled Taylor expansion method to first order [17–19], obtaining

$$|\psi(t+dt)\rangle = |\psi(t)\rangle - iH|\psi(t)\rangle dt, \tag{7}$$

where $|\psi(t)\rangle$ is the wave function at time t, dt is the time step, and H is the Hamiltonian dictating the time evolution. The simulation time t and timesteps dt are in units of the bandwidth W of the system. We choose $dt = 0.01$, but we have separately checked that our results hold also for smaller time steps. In our simulations we initialize the system from the same real-space Gaussian wave packet. In order to ensure that the wave packets do not reach the system boundary, we limit the total number of time steps used for a simulation to $t_{\max} = L/(a \cdot dt \cdot v_{\mathrm{drift}})$, with L the system size, a the lattice constant and v_{drift} the drift velocity of the wave packet for low disorder. Above the localization transition, $t_{\max}$ is not shortened in order to record instances of 'teleportation' of the drift center of the maximally amplified wave packet, which contribute to the average velocity.

The method is based on the following expression for the matrix exponential

$$\mathrm{e}^{-itH} = \lim_{N\to\infty} \left(I - \frac{itH}{N}\right)^N, \tag{8}$$

where N is the number of time steps. Fixing the time step ($dt = t/N$) and the number of steps (N) we get an approximation for the time evolution operator as:

$$\mathrm{e}^{-itH} \approx (I - idtH)^N. \tag{9}$$

The error introduced at each subsequent time step can be estimated using the errors calculated for Taylor polynomials of the first order as [18, 19]:

$$\delta = \left\| \mathrm{e}^{-idtH} - I + idtH \right\| \leq \frac{dt^2 \|H\|^2}{2} \frac{1}{1 - \frac{dt\|H\|}{3}}, \tag{10}$$

where $\|\cdot\|$ is any well defined matrix norm, for simplicity we use the spectral norm. For normalized Hamiltonians $\|H\| = 1$ and $dt \leq 1$, the error introduced at each time step is $\delta \leq 3dt^2/4$.

C Analytical form of wave packet center of mass drift velocity

In this section we provide details concerning the derivation of Eq. (2).
Starting from $\partial_t[\langle\psi|\hat{x}|\psi\rangle/\langle\psi|\psi\rangle]$ and using the quotient rule,

$$\partial_t[\langle\psi|\hat{x}|\psi\rangle/\langle\psi|\psi\rangle] = \frac{\partial_t[\langle\psi|\hat{x}|\psi\rangle]\cdot\langle\psi|\psi\rangle}{\langle\psi|\psi\rangle^2} - \frac{\langle\psi|\hat{x}|\psi\rangle\cdot\partial_t[\langle\psi|\psi\rangle]}{\langle\psi|\psi\rangle^2}. \tag{11}$$

We first calculate $\partial_t[\langle\psi|\psi\rangle]$. Using

$$|\psi\rangle = e^{-itH}|0\rangle, \tag{12}$$

$$\langle\psi| = \langle 0|e^{itH^\dagger}, \tag{13}$$

we find

$$\partial_t[\langle\psi|\psi\rangle] = \partial_t[\langle 0|e^{itH^\dagger}e^{-itH}|0\rangle] \tag{14}$$

$$= \langle 0|iH^\dagger e^{itH^\dagger}e^{-itH}|0\rangle + \langle 0|e^{itH^\dagger}(-iH)e^{-itH}|0\rangle \tag{15}$$

$$= i\langle\psi|H^\dagger - H|\psi\rangle. \tag{16}$$

Similarly for $\partial_t[\langle\psi|\hat{x}|\psi\rangle]$,

$$\partial_t[\langle\psi|\hat{x}|\psi\rangle] = \partial_t[\langle 0|e^{itH^\dagger}\hat{x}e^{-itH}|0\rangle] \tag{17}$$

$$= \langle 0|iH^\dagger e^{itH^\dagger}\hat{x}e^{-itH}|0\rangle + \langle 0|e^{itH^\dagger}\hat{x}(-iH)e^{-itH}|0\rangle \tag{18}$$

$$= i\langle\psi|H^\dagger\hat{x} - \hat{x}H|\psi\rangle. \tag{19}$$

Substituting these expressions into Eq. (11), we obtain

$$\partial_t[\langle\psi|\hat{x}|\psi\rangle/\langle\psi|\psi\rangle] = i\langle\psi|H^\dagger\hat{x} - \hat{x}H|\psi\rangle - i\langle\psi|\hat{x}|\psi\rangle\langle\psi|H^\dagger - H|\psi\rangle, \tag{20}$$

where we use that $\langle\psi|\psi\rangle = 1$.
We then express Eq. (20) in terms of the current operator $J = i[H,\hat{x}]$. Using

$$\mathrm{Re}(J) = -\mathrm{Im}([H,\hat{x}]) = \frac{i}{2}(\langle\psi|[H,\hat{x}]|\psi\rangle - \langle\psi|[\hat{x},H^\dagger]|\psi\rangle) \tag{21}$$

$$= \langle\psi|(H\hat{x} - \hat{x}H - \hat{x}H^\dagger + H^\dagger\hat{x})|\psi\rangle, \tag{22}$$

we obtain

$$\partial_t[\langle\psi|\hat{x}|\psi\rangle/\langle\psi|\psi\rangle] = \mathrm{Re}(J) + \frac{i}{2}\langle\psi|(H^\dagger\hat{x} - \hat{x}H - H\hat{x} + \hat{x}H^\dagger)|\psi\rangle - i\langle\psi|\hat{x}|\psi\rangle\langle\psi|H^\dagger - H|\rangle. \tag{23}$$

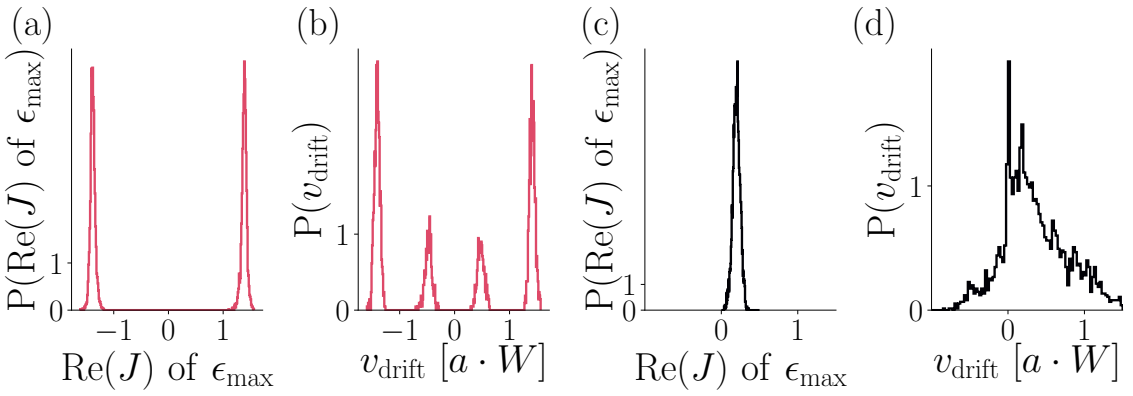


Figure 4: Multimodal and bimodal distributions of $\mathrm{Re}(J)$ and v_{drift} of the Hamiltonians ~~H_8~~H_∞ [Eq. (4)] and H_{HN} [Eq. (1)]. For the Hatano-Nelson Hamiltonian (1), random disorder terms $U_{i,j}$ with $i \in \{0,1,2\}$ and j the site number are sampled from distributions with standard deviation δ_i. We set $\delta_0 = \delta_1 = \delta_2 \equiv \delta$. For the Hamiltonian H_∞ (4), random disorder terms $U_{i,j}$ with $i \in \{0,1,2,3,4\}$ and j the site number are sampled from distributions with standard deviation δ_i. We set $\delta_0 = \delta_1 = \delta_2 = \delta_3 = \delta_4 \equiv \delta$. (a), (c) distributions of $\mathrm{Re}(J)$ of the maximally amplified state and v_{drift} of the ~~H_8~~H_∞ model at $\phi = 0$ and $\delta = 0.1$. (b), (d) distributions of $\mathrm{Re}(J)$ of the maximally amplified state and v_{drift} of the H_{HN} model at $\phi = 0$ and $\delta = 0.8$. Plot details in App. A.

We rewrite the second and third terms of Eq. (23) as

$$\partial_t[\langle\psi|\hat{x}|\psi\rangle/\langle\psi|\psi\rangle] = \mathrm{Re}(J) + \frac{i}{2}\langle\psi|(H^\dagger\hat{x} - \hat{x}H - H\hat{x} + \hat{x}H^\dagger + 2\langle\psi|\hat{x}|\psi\rangle(H - H^\dagger))|\psi\rangle \tag{24}$$

$$= \mathrm{Re}(J) + \frac{i}{2}\langle\psi|((H - H^\dagger)(2\langle\psi|\hat{x}|\psi\rangle - \hat{x}) - \hat{x}(H - H^\dagger))|\psi\rangle \tag{25}$$

$$= \mathrm{Re}(J) + \frac{i}{2}\langle\psi|((H - H^\dagger)(\langle\psi|\hat{x}|\psi\rangle - \hat{x}) - (\langle\psi|\hat{x}|\psi\rangle - \hat{x})(H - H^\dagger))|\psi\rangle \tag{26}$$

$$= \mathrm{Re}(J) + \frac{i}{2}\langle\psi|\{(H - H^\dagger), (\langle\psi|\hat{x}|\psi\rangle - \hat{x})\}|\psi\rangle. \tag{27}$$

D Multimodal behavior

Here we discuss the shape of the distributions of $\mathrm{Re}(J)$ and v_{drift} of both the ~~H_8~~H_∞ [Eq. (4)] and H_{HN} [Eq. (1)] models around the transition point.

For ~~H_8~~H_∞, the distribution of $\mathrm{Re}(J)$ of the maximally amplified eigenstate is bimodal [Fig. 4 (a)]. The distribution of v_{drift} is multimodal [Fig. 4 (c)]. The multimodality arises from the disorder nontrivially shifting eigenvalues of ~~H_8~~H_∞ in the complex plane, creating two bimodal distributions for v_{drift}, one on each side of the transition in ϕ. The same multimodal behavior is seen in $\mathrm{Re}(J)$ of the maximally amplified eigenstate when using biorthogonal expectation values to calculate J [App. E, Fig. 5 (a)].

For H_{HN}, $\mathrm{Re}(J)$ does not exhibit a transition [Fig. 3 (d)], and its distribution close to the v_{drift} transition is centered around a small but finite value [Fig. 4 (b)]. The scaling of

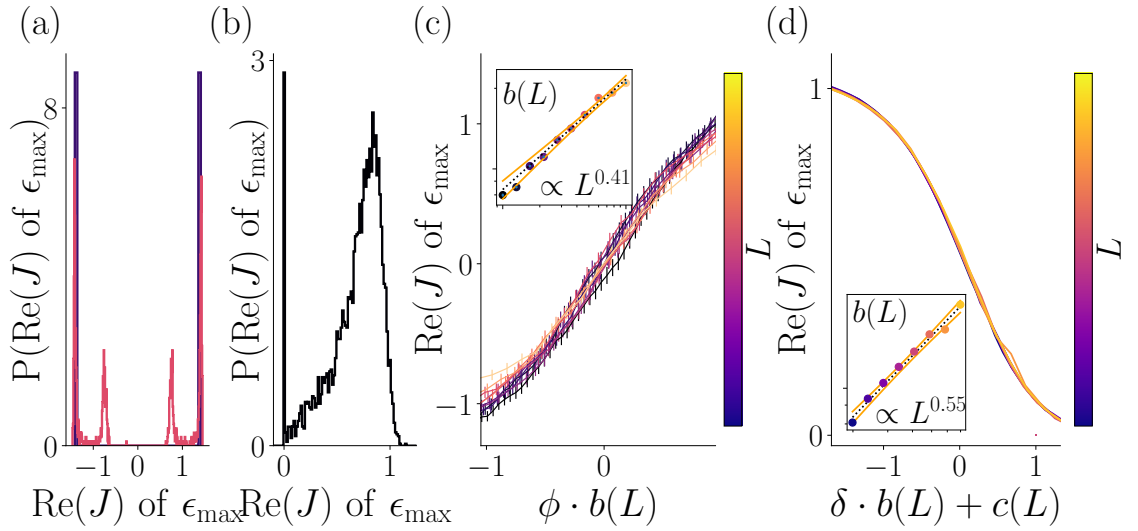


Figure 5: $\mathrm{Re}(J)$ results using biorthogonal expectation values for Hamiltonians ~~H_8~~ H_∞ [Eq. (4)] and H_{HN} [Eq. (1)]. For the Hatano-Nelson Hamiltonian (1), random disorder terms $U_{i,j}$ with $i \in \{0,1,2\}$ and j the site number are sampled from distributions with standard deviation δ_i. We set $\delta_0 = \delta_1 = \delta_2 \equiv \delta$. For the Hamiltonian H_∞ (4), random disorder terms $U_{i,j}$ with $i \in \{0,1,2,3,4\}$ and j the site number are sampled from distributions with standard deviation δ_i. We set $\delta_0 = \delta_1 = \delta_2 = \delta_3 = \delta_4 \equiv \delta$. (a) distributions of $\mathrm{Re}(J)$ of the maximally amplified state of the ~~H_8~~ H_∞ model at $\phi = 0$ and $\delta = 0.1$ and $\delta = 0.01$ in units of the bandwidth W. (b) distribution of $\mathrm{Re}(J)$ of the maximally amplified state of the H_{HN} model at $\delta = 0.8$. (c) Rescaled $\mathrm{Re}(J)$ of the maximally amplified state of the ~~H_8~~ H_∞ model for $\delta = 0.1$. (d) Rescaled $\mathrm{Re}(J)$ of the maximally amplified state of the H_{HN} model for $\delta = 0.1$. Insets of (c) and (d): scaling functions of the slope at the transition and the 95% confidence interval. Plot details in App. A.

v_{drift} of the Hatano-Nelson model H_{HN} does not exactly follow $\sqrt{t_{\max}}$ [Table 1 and Fig. 3 (d)] but a bimodal distribution is still observed close to the transition [Fig. 4 (d)]. Close to the transition point, the distribution of v_{drift} has two peaks, with one broad peak centered around a finite value, and the other delta function peak around 0. The v_{drift} around 0 originates from disorder configurations that result in localization, and the v_{drift} with finite velocity originates from disorder configurations where propagation is still possible.

E Biorthogonal expectation value

In the results of the manuscript, we calculated $\mathrm{Re}(J)$ of the state m as $\mathrm{Re}\left(\langle\psi_m|J|\psi_m\rangle\right)$ such that $\langle\psi_m| = |\psi_m\rangle^\dagger$. In this section we calculate $\mathrm{Re}(J)$ of state m as $\mathrm{Re}\left(\langle\psi_m|J|\psi_m\rangle\right)$ such that $\langle\psi_m| = |\psi_m\rangle^{-1}$, that is to say $\langle\psi_m|$ is the m-th left eigenstate and $|\psi_m\rangle$ is the m-th right eigenstate. We refer to this $\mathrm{Re}\left(\langle\psi_m|J|\psi_m\rangle\right)$ as the biorthogonal expectation value of $\mathrm{Re}(J)$. The behavior of $\mathrm{Re}(J)$ is significantly impacted by this change in expectation value, as shown in Fig. 5.

For the ~~H_8~~ H_∞ model, similarly to Fig. 4 for low disorder ($\delta = 0.01$) the distribution of $\mathrm{Re}(J)$ of the maximally amplified eigenstate is bimodal [Fig. 5 (a)]. At finite disorder, the distribution becomes multimodal, similar to v_{drift} [Fig. 4 (c)]. The scaling parameter at the transition of the biorthogonally projected $\mathrm{Re}(J)$ scales as $L^{0.44\pm0.01}$ [Fig. 5 (c)].

The Hatano-Nelson model H_{HN} also exhibits bimodal behavior [Fig. 5 (b)], similarly

Parameter	value
$t_{x,+}$	1
$t'_{x,+}$	0
$t_{x,-}$	0.8
$t'_{x,-}$	0
$t_{y,+}$	0
$t'_{y,+}$	0
$t_{y,-}$	0
$t'_{y,-}$	1

Table 2: Parameters used for simulating Hamiltonian Eq. (28).

to the distribution of v_{drift} [Fig. 4 (d)]. Close to the transition point, the distribution of $\text{Re}(J)$ has two peaks, with one broad peak around the low-disorder $\text{Re}(J)$ value 1.1, and the other delta function peak around the high disorder $\text{Re}(J)$ value 0. In the biorthogonal case, the $\text{Re}(J)$ of the H_{HN} displays a phase transition. The scaling of the transition width is found to scale close to $\sqrt{L}$, as $L^{0.55\pm0.02}$ [Fig. 5 (d)], similarly to the ~~H_8~~ H_∞ case.

The $\text{Re}(J)$ calculated using the biorthogonal expectation value appears to follow the behavior of v_{drift} more closely, but we do not have an argument as to why this would be the case.

F Results in two dimensions

We consider the following two-dimensional non-Hermitian model:

$$\begin{aligned} H_N = &\sum_{d=1}^{N}\sum_{j}^{L_d} U_{0,j}|x_{d,j}\rangle\langle x_{d,j}| + \left(t_{x_d,+} + it'_{x_d,+}U_{1,d,j}\right)|x_{d,j+1}\rangle\langle x_{d,j}| \\ &+ \left(t_{x_d,-} + it'_{x_d,-} + U_{2,d,j}\right)|x_{d,j}\rangle\langle x_{d,j+1}|, \end{aligned} \tag{28}$$

where the sum runs over all the lattice sites j and the spatial dimensions d of a N-dimensional system with $L/a = \frac{1}{a}\sum_d^N L_d$ sites, with a the lattice constant. x_d corresponds to the spatial coordinate in dimension d, and $U_{n,d,j}$ are random disorder terms sampled from a distribution $\delta_{n,d}$. We choose $N = 2$.

The parameters we use in simulating this model are found in Table 2, and yield the spectrum shown in Fig. 6 (a)-(b). Wave packets are initialized at the center of the periodically wrapped lattice ($x_0 = [0, 0]$), with a width one tenth of the width of the lattice ($\sigma = L/10$) and with the same initial velocity ($k_x = \pi/2, k_y = 0$) for all simulations.

We fit the function $a\tanh(b\delta + c)$ to the localization transition of v_{drift} as a function of δ, and extract $b(t_{\max})$. $b(t_{\max})$ scales as $t_{\max}^{0.63\pm0.05}$ [Fig. 6 (d)]~~. The critical exponent of 2D non-Hermitian dynamic systems approaches $\nu = 0.5$. However it~~ , which differs from the 1D critical exponents presented in the main text [Table 1]. It is not possible from these results to say whether the critical exponent of non-Hermitian systems is dimension-dependent or not.

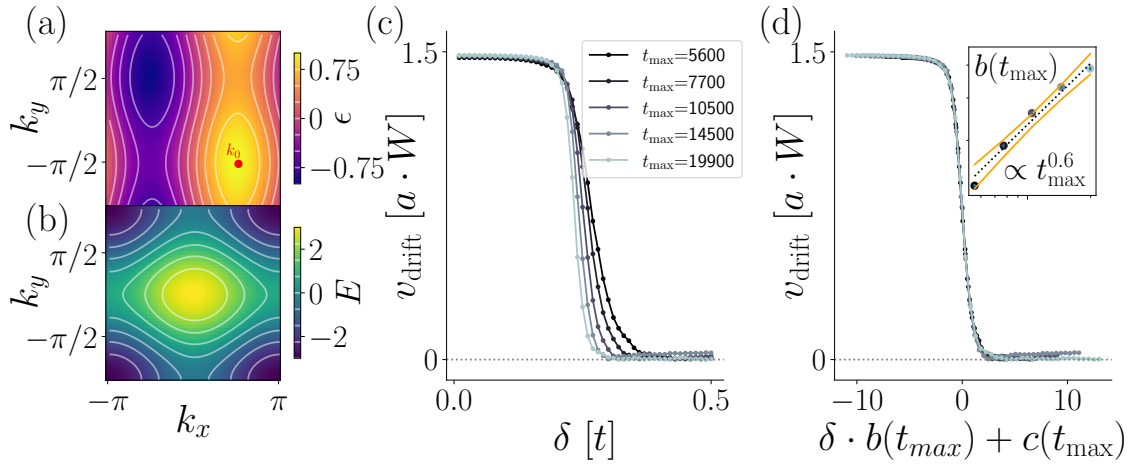


Figure 6: Finite-time scaling of two-dimensional non-Hermitian model Eq. (28) with parameters 2. (a)-(b) The Brillouin zone of (a) the imaginary part of the energy ϵ and (b) the real part of the energy. (c) The unscaled localization transition of v_{drift} as a function of ~~δ~~disorder. Random disorder terms $U_{i,d,j}$ with $i \in \{0,1,2\}$, $d \in \{x,y\}$ and j the site number (28) are sampled from distributions that all have a standard deviation of $\delta_0 = \delta_1 = \delta_2 \equiv \delta$. (d) The rescaled curves of (c). Inset: scaling of the sharpness of the transition. Plot details in App. A.

G Unscaled results

The results for $\text{Re}(J)$ at ϵ_{max} and v_{drift} shown in Fig. 2 and Fig. 3 are rescaled by the scaling variables b (and c in the case of v_{drift}). Fig. 7 contains the unscaled data used to obtain Fig. 2 and Fig. 3, as well as the rescaled data for comparison.

We do not show rescaling of the $\text{Re}(J)$ at ϵ_{max} curves of Fig. 7 (c1), since they do not exhibit scaling.

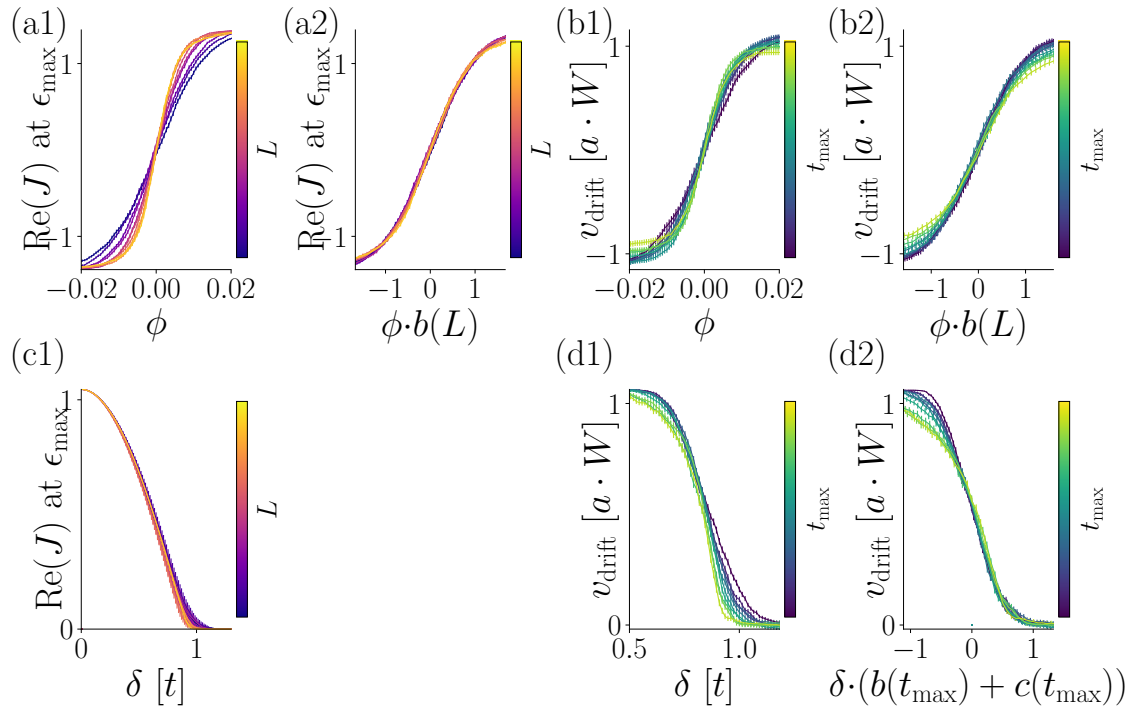


Figure 7: Unscaled (a1,b1,c1,d1) and rescaled (a2,b2,d2) $\mathrm{Re}(J)$ at the point of maximal amplification $\epsilon_{\max}$ and v_{drift} at the transition point, as a function of ϕ the tilt angle of the spectrum or as a function of disorder δ. For the Hatano-Nelson Hamiltonian (1), random disorder terms $U_{i,j}$ with $i \in \{0,1,2\}$ and j the site number are sampled from distributions with standard deviation δ_i. We set $\delta_0 = \delta_1 = \delta_2 \equiv \delta$. For the Hamiltonian H_∞ (4), random disorder terms $U_{i,j}$ with $i \in \{0,1,2,3,4\}$ and j the site number are sampled from distributions with standard deviation δ_i. We set $\delta_0 = \delta_1 = \delta_2 = \delta_3 = \delta_4 \equiv \delta$. (a1)-(b2) Results for the ~~H_8~~ H_∞ model Eq. (4) used in Fig. 2. (c1)-(d2) Results for the H_{HN} model Eq. (1) used in Fig. 3. Plot details in App. A.